What are reaction time indices of automatic imitation measuring?

Ramsey, Richard

Consciousness and Cognition

DOI:
10.1016/j.concog.2018.08.006

Published: 01/10/2018

Peer reviewed version

Journal: Consciousness and Cognition

Title: What are reaction time indices of automatic imitation measuring?

Richard Ramsey

Wales Institute for Cognitive Neuroscience, School of Psychology, Bangor University, Bangor, Gwynedd, Wales, LL57 2AS, United Kingdom

Correspondence: r.ramsey@bangor.ac.uk

Word count:

Key words: automatic imitation; reaction time measures; validity; domain-specificity; social cognition.

Abstract

Humans unintentionally copy other people's gestures, postures and speech patterns. This behaviour has been termed 'automatic imitation', and is thought to play a crucial role in the development of social relationships by building affiliation and rapport between interaction partners. The current paper critically evaluates reaction time measures of automatic imitation and demonstrates that it is currently unclear what reaction time indices measure, due to a lack of research assessing dimensions of validity and domain-specificity. Based on this assessment of the literature, three suggestions are made for future research. First, studies that assess the validity of reaction time paradigms as an index of automatic imitation are required. That is, the extent that processes involved in reaction time measures of automatic imitation also operate in real-world copying behaviours needs firmly establishing. Second, the extent that component processes of automatic imitation rely on domain-specific, domain-general or a combination of both types of cognitive system needs establishing. Moreover, theories of imitation that move beyond neat divisions between domain-specific and domain-general systems should be prioritised. Third, a case is made for why methodological rigour requires far more emphasis in order to support the development of a cumulative science of automatic imitation. In sum, by providing a critical review of the current literature, as well as suggesting new directions for research questions, theories and methods, this paper aims to stimulate discussion regarding optimal approaches to studying automatic imitation. Moreover, the analysis and recommendations presented here generalise to other domains of social cognition, such as gaze perception and perspective taking, which also rely on similar reaction time paradigms.

1. Introduction

During social interactions, we unintentionally copy other people's actions. For example, we may copy an interaction partner's body posture, gestures and speech patterns without intending to do so. Such automatic imitation behaviour has been argued to serve a central social function by building affiliation, rapport and liking between individuals (Chartrand & Lakin, 2013).

Given the ubiquitous influence of imitation on our daily lives, it is not surprising that researchers in diverse sub-disciplines of psychology and neuroscience have studied imitation as a means to better understand human social behaviour. Across these research domains, a variety of methods have been adopted to study imitation, which vary in ecological validity (Figure 1). In developmental and social psychology, for example, methods typically involve observing live interactions and recording overt copying behaviours (Chartrand & Larkin, 2013; Ray & Heyes, 2011). The advantage of observing live social interactions is the relatively straightforward sense in which copying behaviours relate to real-life imitation behaviour. By contrast, in cognitive psychology, reaction time measures of automatic imitation have been designed based on stimulus-response compatibility (SRC) paradigms (Brass et al., 2000; Stürmer et al., 2000). SRC paradigms have the advantage of being amenable to study in controlled laboratory environments, but the disadvantage of reduced ecological validity to real-life copying behaviour. Drawing links between reaction time measures and real-world imitation behaviour is therefore less direct than observational approaches and requires further validation.

The validity of a measure, along with its reliability, are important concerns for any scientific method of measurement: Are the tools used to study a particular phenomenon consistent across time (are they reliable?) and do they measure what they intend to measure (are they valid?)? Here, I address these questions with regard to automatic imitation research

with a particular focus on reaction time measurements. In addition, I also review evidence that speaks to the specificity or generality of the component processes that make up automatic imitation. Like many complex cognitive processes, assessing evidence for domain-specificity is important when interpreting SRC measures of automatic imitation. For example, SRC measures could be a valid measure of real-world copying tendencies but operate on domain-general cognitive and neural architectures for prioritising and guiding behaviour. As such, in addition to reliability and validity, I outline how it is also important to concurrently consider domain-specificity when interpreting component processes of SRC measures of automatic imitation.

The paper is structured in four parts. First, I provide a short review of reaction time tasks that are typically employed in cognitive psychology to study imitation (for a detailed review, see Heyes, 2011). Second, I review evidence that assesses the reliability and validity of these RT measures. I show that we currently know very little about the reliability or validity of RT measures of automatic imitation and this presents a major obstacle to promoting a progressive research programme (Lakatos, 1970; Meehl, 1990). Third, I consider the extent to which component processes of automatic imitation rely on domain-specific and domain-general cognitive and neural mechanisms. I show that there is evidence for a strong domain-general contribution to some component processes of imitation, which involve the control of automatic imitation, which is contrary to a dominant view in the literature. In the final part of the paper, I present an outlook for future research using SRC measures of imitation. I propose that future work should focus on establishing validity of the measure, as well as the relative domain-specificity of component processes. In addition, I propose several ways that models of imitation might be revised to more clearly account for interactions between component processes, as well as how increased methodological rigour should help to advance this research programme in a progressive and cumulative manner.

The overarching aims of this paper are to open a discussion regarding optimal methods to study automatic imitation using reaction time measures and to highlight several useful theoretical and methodological approaches that are currently being overlooked. Importantly, the suggestions for future research generalise to other domains of social cognition such as gaze perception and perspective taking, which also rely on SRC paradigms. Overall, I argue that SRC paradigms have the potential to make a valuable contribution to understanding mechanisms associated with our pervasive tendency to copy others, as well as other dimensions of social cognition, but that the current use of such paradigms requires revision.

2. Reaction time measures of automatic imitation

2.1 Key features of the task

SRC paradigms were initially developed by two research groups with an aim to index processes associated with automatic imitation (Brass et al., 2000; Stürmer et al., 2000). SRC paradigms are quite distinct from other measures of automatic imitation, which are used in different psychology sub-disciplines (and typically labelled "mimicry"). For example, in social psychology, imitation is typically measured by observing face-to-face interactions between a participant and a confederate (Chartrand & Lakin, 2013). During these live interactions, the participant is instructed to take part in a task that is independent to imitation (e.g., evaluate photographs), and the number of copying behaviours made by the participant are recorded (Chartrand & Bargh, 1999; Figure 1). As such, the measured behaviour is fully expressed automatic copying during a live social interaction. In other words, the task appears to have high face-validity, even though the reliability of such measures may be quite low (Genschow et al., 2017).

In contrast, SRC measures of automatic imitation are more removed from live imitative behaviour and typically involve computerised testing procedures in a laboratory setting and the measurement of reaction times (Brass et al., 2000; Heyes, 2011). As such, SRC measures do not measure overt copying behaviours. Instead, it is suggested that the impulse to copy is captured by a reaction time signature. In one version of the SRC imitation paradigm, which involves a choice reaction time paradigm, participants respond to an imperative cue, such as a number, whilst also observing an action. For example, participants may be instructed to lift their index finger if a number 1 is presented and lift their middle finger if a number 2 is presented (Figure 1). Concurrent with this instruction, participants also observe a finger lift that can either be compatible or incompatible with the participant's finger lift[1]. Reaction times are longer in the incompatible than compatible condition. Therefore, in a similar manner to other SRC paradigms, task-irrelevant features influence reaction time performance.

In a different version of the task, which involves a simple reaction time paradigm, actions are pre-specified before the trial begins. At the start of every trial, the action to be performed when the imperative cue appears is pre-selected (i.e., always lift your index finger when you see a finger movement). Again, the observed movement can be compatible or incompatible with the performed movement. Under these conditions, compatibility effects still occur to reaction times (Brass, Bekkering & Prinz , 2001; Heyes et al., 2005). Evidence for compatibility effects in simple reaction time paradigms suggest that the act of representing an action can facilitate or interfere with pre-specified movement intentions and that interference effects are not solely tied to situations where a choice between two actions has to be made on every trial. Together, the results from simple and choice SRC paradigms

[1] Variants of this method have used a hand opening and closing instead of finger movements, but the logic of the SRC approach is identical (e.g., Press et al., 2005; Wang et al., 2011).

demonstrate that a task-irrelevant feature, the appearance of a moving body part, produces reaction time facilitation in the compatible condition and interference in the incompatible condition.

Importantly, the difference in reaction time between compatible and incompatible conditions or "compatibility effect" cannot be reduced to differences in spatial compatibility, such that up-down, left-right movements of a limb could produce the same effects (for a review, see Heyes, 2011). Instead, SRC measures of automatic imitation relate to the topographical or configural features of the observed body in relation to one's own bodily features and motor intention (Heyes, 2011). Example features of SRC measures of automatic imitation include interference from different types of effectors and movements. For instance, effector compatibility refers to whether the identity of the observed effector (e.g., finger, arm, mouth) is compatible with one's own effector, whereas movement compatibility refers to whether the type of movement observed (e.g., tapping, waving, smiling) is compatible with one's own movement. For both effector and movement compatibility, the observed action is represented and in the compatible case it facilitates task performance and in the incompatible case it interferes with task performance.

At least two processes are engaged for both compatible and incompatible conditions: action representation and action control (Figure 2). The distinction between representation and control follows directly from work in semantic cognition, which distinguishes between representation and control in support of understanding meaning (Jefferies, 2013; Lambon Ralph et al., 2017). In the domain of semantic cognition, the representational system refers to knowledge of concepts, which are a product of higher-order relationships between sensory, motor, linguistic and affective sources of input. For example, the meaning of a concept such as "dog", has been argued to be a product of associated sensory, motor, linguistic and affective properties, which are distributed across multiple input modalities. By contrast, the

control system manipulates the representational system as a function of the relevant context. For example, when playing football, a dominant response may be to activate a kicking action on seeing a football. But, when taking a throw-in, this dominant response needs to be controlled in order to make use of the hands. Indeed, in everyday life, the dominant or pre-potent association may not be appropriate. Instead, a non-dominant aspect of knowledge or motor responses may need to be retrieved in a more controlled way by prioritising some features or actions over alternatives. As such, the interaction between representation and control systems, as well as their independent contributions, are central features to understanding sematic cognition (Lambon Ralph et al., 2017).

The logic of two principal interacting systems can be applied to the context of automatic imitation. First, observed actions are detected and represented. Such action representation is likely to involve both perceptual and motor representations, which span visual and sensorimotor cortices (Caspers et al., 2010). Further, but more speculatively, the process of action representation could also involve a "matching" process, which compares the observed movement to one's own motor state and likely engages the mirror neuron system (Rizzolatti & Craighero, 2004; Iacoboni, 2009). Second, a selection or control mechanism is required to implement a chosen action according to task instructions (Brass et al., 2009; Cross et al., 2013). In the compatible case, the selection process is assisted by virtue of observing the same action feature, whereas in the incompatible case the selection process is more challenging as the observed action needs to be suppressed and the correct action prioritised, in order for the task to be completed correctly. Of course, in simple SRC paradigms, the same level of demand is not placed on action selection processes as the action is pre-specified. Instead, the previously selected action needs to be executed whilst observing ongoing actions. As a consequence of these SRC effects, it has been argued that the compatibility effect is an index of the additional cognitive resources required to suppress the incompatible observed

action and enforce one's own intended action (Brass & Heyes, 2005; Brass et al., 2009; Heyes, 2011). In other words, the compatibility effect partly indexes control over the spontaneous urge to imitate, such that incompatible actions are suppressed and one's motor intentions are prioritised.

A related, but more cognitively elaborate interpretation, suggests that the congruency effect indexes a process of self-other distinction (Brass & Heyes, 2005; Brass et al., 2009). That is, in order to perform the task, one has to be able to distinguish between one's own motor intention from the motor intentions of others. It has been suggested that such a self-other distinction relies on key nodes within the theory of mind network including medial prefrontal cortex and right temporoparietal junction (Brass & Heyes, 2005; Brass et al., 2009). The theory of mind network has been consistently associated with a range of social cognition processes (van Overwalle, 2009). Further, right temporoparietal junction has been implicated in distinguishing actions of self and other (Farrer & Frith, 2002; Ruby & Decety, 2001). Hence, according to one set of related views, the congruency effect indexes a cognitive mechanism that is shared with processing in other contexts that also require a self-other distinction. Consequently, a self-other distinction underpinned by right temporoparietal junction has been argued to play a broader role in social cognition more generally, including autism, empathy and theory of mind (de Guzman et al., 2016; Sowden & Shah, 2014; Spengler et al., 2009).

Representation and control in automatic imitation are likely to rely on a set of distributed neural circuits (Figure 2). Furthermore, in the context of automatic imitation, evidence to date that identifies the neural substrates supporting action representation appears more settled than that underpinning action selection and control. For example, person perception processes, which detect others through representation of faces, bodies and biological motion, span the ventral visual stream and occipitotemporal cortices (Kanwisher,

2010). In addition, action representations evoked by observing others span frontoparietal cortices, which include the mirror neuron system, as well as occipitotemporal cortices (Caspers et al., 2010; Molenberghs et al., 2012). In contrast, in terms of action selection and control, three proposals have been put forward, which are not necessarily mutually exclusive. First, the multiple demand network has been argued to play a domain-general role in the control of automatic imitation (Cross et al., 2013; Darda, Butler & Ramsey, 2018). The multiple demand network spans dorsolateral and dorsomedial frontoparietal cortex and has been associated with a range of executive functions, such as inhibition and working memory (Duncan, 2010). A second type of account relies on the theory of mind network to provide forms of "social control" over imitative tendencies in a domain-specific manner (Brass & Heyes, 2005; Brass et al., 2009; Wang & Hamilton, 2012). A third account suggests that the mirror neuron system may contribute to action selection and such a selection mechanism may be especially relevant in imitative contexts (Hickok & Hauser, 2010; Hickok, 2013). These accounts of action representation and control are illustrated in Figure 2 and it is clear that further research is required to distinguish between these accounts, especially in terms of control.

Ultimately, and irrespective of the cognitive interpretation placed on the congruency effect, the SRC automatic imitation task has been argued to be a lab equivalent of overt copying behaviour measures (Heyes, 2011). In overt copying behaviour, automatic imitation is fully expressed as a copying action, whereas the SRC measure is argued to index a covert signature of the unexpressed tendency to automatically imitate. As such, the possible utility of such a task is rather enticing, as it would enable questions of social cognition to be probed with the experimental control afforded by laboratory conditions typical of SRC measures. With the obvious interdisciplinary appeal of linking SRC tasks to overt copying behaviours, it is somewhat surprising that although many papers have been published that cite the original

work (i.e., according to Google Scholar in July 2018, Brass et al., 2000 has been cited over 650 times), there has been very little empirical work that has directly assessed the reliability or validity of the task, which is the focus of the next section.

2.2. Reliability and validity

Reliability and validity are key concerns for any type of scientific research. Although the SRC measure of automatic imitation has been used as a tool to study the intersection between social and cognitive psychology, as well as cognitive neuroscience, it is surprising that very few studies have investigated its reliability or validity. To date, only one study has systematically studied reliability and validity in the SRC task (Genschow et al., 2017). Genschow and colleagues (2017) demonstrated that the reaction time compatibility effect had a large effect size (Cohen's d = 1.56) and high levels of split-half reliability (Spearman-Brown coefficient =*.86*). That is, there was relatively good correspondence between task performance during odd and even trials. By contrast, Genschow and colleagues (2017) showed that a measure of overt copying behaviour (rather than reaction times) had low reliability and overt copying behaviour did not correlate with performance on the SRC paradigm. Of course, low reliability in one task makes it hard to interpret a lack of correlation between variables (Cronbach, 1990). Hence, to date, no robust evidence currently exists either for or against the validity of SRC tasks as a laboratory equivalent of overt copying of other people's behaviour. As such, it seems necessary to establish the validity of the task, in terms of its ability to index particular types of imitative process.

As recently summarised by Gray (2017), at least three types of validity would be important to establish with respect to the SRC task: construct, convergent and divergent validity. In essence, it is important to establish what SRC paradigms are measuring, as well as what these measurements relate to and do not relate to. These fundamental relationships need

specifying before researchers can use the SRC paradigm to draw strong inferences regarding imitation-related processes and other social or non-social parameters. The importance of each type of validity in the context of automatic imitation research is outlined below.

Construct validity

Construct validity is the degree to which a test measures what it claims to be measuring (Cronbach & Meehl, 1955). Although commonly applied to psychometric studies in psychology, evidence that reaction time measures of automatic imitation relate to other forms of imitative abilities would provide a valuable source of construct validity. If the SRC paradigm indexes the control of the urge to imitate, then a relationship should exist between other measures that involve the tendency to imitate. Establishing construct validity would provide justification that the SRC paradigm can be used in a manner that indexes processes relating to imitation. Currently, with only one study addressing this issue directly to my knowledge (Genschow et al., 2017), the relationship between reaction time and observational measures of automatic imitation is unclear. As such, before concluding that SRC measures provide a lab equivalent of live copying behaviour (Heyes, 2011), more research is required to validate which processes associated with overt copying behaviour, if any, relate to reaction time measures of imitation. To do so, comparing SRC measures with overt copying behaviours during live social interactions (e.g., Chartrand & Lakin, 2013), as well as using virtual reality (Pan & Hamilton, 2018), would be beneficial.

Even if reaction time and observational indices of imitation are not measuring the same thing, they could be indexing different component processes of a broader phenomena – automatic imitation (Genschow et al., 2017). Therefore, it is possible that both measures could be interesting to study in their own right, in order to understand mechanisms underpinning imitation behaviour, while not sharing any direct relationship to each other.

Firm support for relationships between distinct measures of imitation will only be possible, however, with dedicated empirical investigation.

Convergent validity

Convergent validity is the degree to which two measures of constructs that theoretically should be related are actually related (e.g., Russell, 1978). If the SRC paradigm indexes processes associated with imitation (Brass & Heyes, 2005; Heyes, 2011) and imitation is associated with social processes more generally (Chartrand & Lakin, 2013), then reaction time measures of imitation should correspond to some degree with related social variables. For example, individual differences in social variables such as empathy (Chartrand & Bargh, 1999), as well as person knowledge (van Baaren et al., 2009), mood (van Baaren et al., 2006; Likowski et al., 2011) and pre-existing rapport (McIntosh, 2006; Stel et al., 2010) have been shown to influence live interaction measures of imitation (for a review, see Chartrand & Lakin, 2013).

The equivalent data for reaction time measures of automatic imitation are not entirely convincing. The largest datasets to date provide no evidence for individual differences in reaction time measures as a function of social dimensions of personality, such as empathy, agreeableness or narcissism (Butler et al., 2015; Genschow et al., 2017). Indeed, a recent meta-analysis shows that the effects of social personality variables on SRC measures of automatic imitation overlap with zero, which is suggestive of no effect (Cracco, Bardi et al., 2018). Therefore, the only studies that show effects of personality on SRC measures either do not replicate when tested in larger samples (e.g., Narcissism; Obhi et al., 2014; Hogeveen & Obhi, 2013) or have only been demonstrated once in small single-experiment samples (e.g., Alexithymia: Sowden et al., 2016).

It is possible that SRC tasks are not particularly well-suited to studying individual differences (Hedge et al., 2017), and experimental work may be more effective at studying how social factors influence automatic imitation. To this end, several recent studies have shown that social factors can influence SRC measures of imitation (Heyes, 2011; Wang & Hamilton, 2012). SRC measures of imitation are increased in situations that promote affiliation through eye contact (Wang et al., 2011), group membership (Gleibs et al., 2016), and facial expressions (Butler et al., 2016), as well as when interacting with more human-like, rather than robotic, agents (Press, 2011; Kilner et al., 2003; Klapper et al., 2014). Other work has shown that when prosocial attitudes are generated (Cook & Bird, 2011; Leighton et al., 2010), or prosocial gestures are signalled (Cracco, Genschow, et al., 2018), imitation increases. These studies suggest that the compatibility effect is modulated by social factors in a broadly corresponding manner to findings from overt measures of behavioural copying (Chartrand & Lakin, 2013). These results have led to theories that suggest an opportunity for social reward can increase imitation (Wang & Hamilton, 2012).

In sum, there is mixed evidence for convergent validity using the SRC automatic imitation task. There is no compelling evidence for convergent validity based on individual difference research, but there is some emerging evidence for convergent validity based on experimental work, which shows that social factors influence overt copying behaviours as well as SRC measures. These initial experimental studies hold promise, but more research is needed with larger sample sizes and clear replications to provide confirmatory evidence. Moreover, it is currently unclear which component processes (i.e., action representation, control or both) are being considered in these studies. Without separating which component processes are under investigation, it could mean that some processes exhibit convergent validity and others do not. Consequently, future work that tests component processes separately for convergent validity would be beneficial. More generally, low reproducibility

rates in psychology, which vary between 25 and 50%, are worth consideration in this context (Open Science Collaboration, 2015). Given the current lack of robust replications that confirm live social interaction measures of imitation and coupled with evidence for low reliability of the effects (Genschow et al., 2017), it would be prudent for the field to perform powerful replications of live copying effects, whilst also comparing such effects to SRC measures.

Divergent validity

Divergent or discriminant validity tests whether measures that are supposed to be unrelated to each other are indeed unrelated (Campbell & Fiske, 1959). If the SRC paradigm indexes imitative control specifically, then such measures should diverge, to some degree, from processes associated with non-imitative control and other non-social processes. Initial evidence for divergent validity has been provided through functional magnetic resonance imaging (fMRI) and behavioural studies. For example, using fMRI, Brass and colleagues (2005) showed a double dissociation in right temporoparietal junction between the SRC measure of automatic imitation and a non-social compatibility task (Brass et al., 2005). The authors suggested that responses in right temporoparietal junction diverge to some degree from a non-social task. However, it is unclear how to interpret such a response, because right temporoparietal junction has been associated with orienting to social and non-social stimuli (Corbetta et al., 2008). In addition, the SRC measure of imitation, which was used by Brass and colleagues (2005), partly reflects spatial compatibility or the tendency to respond faster to a stimulus when it is on the same side of space as the response (e.g. Simon 1969). As such, the most commonly used automatic imitation task has non-social (spatial) features, as well as social (imitative) features. Therefore, differences in right temporoparietal junction could reflect different degrees of non-social control, as much as different degrees of social control.

Behavioural data have also provided some evidence for divergent validity by measuring the imitative component independent from the spatial component. These studies show that in the absence of a spatial component, the imitative component remains above zero, but is 3-4 times smaller than the spatial component (Bertenthal et al., 2006; Boyer et al., 2012; Catmur & Heyes, 2011; Darda et al., 2018; Gowen et al., 2016; Marsh et al., 2016). Additionally, the spatial and imitative components display different timecourses, which further demonstrates that they are dissociable (Catmur & Heyes, 2011). Whilst these behavioural findings show that spatial and imitative components of the task are not identical and diverge to some degree, the lack of robust evidence for construct and convergent validity for the imitative component mean that it is still difficult to interpret as a valid reaction time signature of overt automatic copying behaviour. Primarily, there is an absence of robust evidence for construct and convergent validity for the imitative component. In addition, however, the evidence that does exist is limited to a few studies, which typically do not separate imitative and spatial component processes. Therefore, the small amount of evidence that is available does not convincingly demonstrate construct or convergent validity for the imitative component of SRC measures.

In sum, the extent to which SRC measures are related to imitation and other social abilities and diverge from non-social abilities will be important for future work to more firmly establish, in order to understand what SRC measures of automatic imitation are and are not indexing.

3. Domain-specificity

A further consideration for interpreting SRC measures of automatic imitation is the extent to which SRC measures rely on domain-specific and/or domain-general processes. Domain-

specific processes are tailored to particular tasks or types of information, whereas domain-general processes operate across multiple tasks and types of information (Barrett, 2012; Fodor, 1983). In this section, I consider current evidence for domain-specificity in SRC measures of automatic imitation, as well as the implications for current theories.

A neat division between domain-specific and domain-general processes is likely to be an unrealistic model for many aspects of perception and cognition (Barrett, 2012; Kanwisher, 2010; Michael & D'Ausilio, 2015; Spunt & Adolphs, 2017). Indeed, consideration of the relationship between domain-specific and domain-general systems provides more fertile ground for understanding complex mental processes and also respects the inherent complexity of mind/brain organisation. Further, most cognitive measures are unlikely to be "pure" measures of a process, which necessarily prompts the question of extent. That is, to what extent do domain general and specific systems contribute to imitative processes and how are they related?

Addressing questions of extent has not been a primary focus of SRC imitation studies. For example, studies that have manipulated imitative compatibility independent of spatial compatibility (e.g., Bertenthal et al., 2006; Boyer et al., 2012; Catmur & Heyes, 2011) are unlikely to represent a pure measure of imitation control. This is especially important for theories of imitation control, which suggest that self-other conflict between interaction partners is regulated through engagement of a domain-specific neural network associated with social cognition (Brass et al., 2009). Specifically, Brass and colleagues (2009) propose that right temporoparietal junction and medial prefrontal cortex, which are key nodes in the theory-of-mind network (Frith & Frith, 1999; Saxe & Kanwisher, 2003; van Overwalle, 2009), are also involved in regulating imitative interactions between self and other. Moreover, a self-other distinction underpinned by right temporoparietal junction has been argued to play a broader role in social cognition more generally, including autism, empathy and theory of

mind (de Guzman et al., 2016; Sowden & Shah, 2014; Spengler et al., 2009). However, right temporoparietal junction is also associated with domain-general processes associated with the multiple demand network (Duncan, 2010) and with orienting to social as well as non-social cues (Corbetta et al., 2008). Thus, right temporoparietal junction has been associated with multiple social and non-social functional processes (Schuwerk et al., 2017).

In terms of fMRI evidence, results published to date do not provide convincing evidence for the engagement of a domain-specific neural network during the control of imitation that is specifically tied to social cognition. Some studies have shown engagement of brain regions associated with the theory of mind network (e.g., Brass, Zysset & von Cramon, 2001; Brass et al., 2005), while others showed engagement of the multiple demand network (e.g., Bien et al., 2009; Marsh et al., 2016). The multiple demand network has been shown to operate in a wide variety of tasks that require processes such as inhibition, working memory and attention (Duncan, 2010). Mental operations in the multiple demand network are domain-general, rather than domain-specific, in the sense that they operate across a variety of contexts. Moreover, such domain-general mental operations have been associated with regions of dorsolateral and dorsomedial frontoparietal cortices (Duncan, 2010; Fedorenko et al., 2013), which are largely distinct from the theory of mind network (Frith & Frith, 1999; Saxe & Kanwisher, 2003; Van Overwalle, 2009).

A limiting factor in most prior fMRI studies of automatic imitation, however, is that they have typically tested small sample sizes and have not functionally defined regions of interest, thus reducing functional resolution. Small sample sizes reduce statistical power to detect effects and low functional resolution means that engagement of a region, especially one with such widespread functional properties as right temporoparietal junction, is difficult to interpret in a functional sense. More recent work by Darda and colleagues (2018), however, has taken steps to overcome these limitations by scanning larger sample sizes in a multi-

experiment study (N=28; N=50), and independently localising theory-of-mind and multiple demand networks in individual participants. Moreover, whilst keeping the stimulus (a human hand) and task (respond to number cues) constant, Darda and colleagues (2018) measured imitative compatibility independently from spatial compatibility and showed that imitative compatibility effects only recruit the multiple demand network and not the ToM network. More specifically, imitative compatibility effects were identified in parts of inferior and superior parietal cortex, as well as inferior and middle frontal gyri, which were independently localised using a task that engages the multiple demand network (Darda et al., 2018). In contrast, in regions of the theory of mind network, even trends in the expected direction were not observed. The evidence for domain-specificity for the control of automatic imitation from fMRI studies, therefore, remains unconvincing to date.

These results have implications for neural and cognitive theories of automatic imitation. The results are consistent with models of imitation that feature a generalist view of control, which relies on the multiple demand network rather than the theory of mind network. Indeed, evidence that control in imitation relies on the same system that is engaged in a wide-range of executive functions such as inhibition and working memory (Darda et al., 2018), may suggest that the same selection mechanism operates the social and non-social contexts. For instance, the control mechanism could be the same whether selecting between social (fingers, arm and mouth movements) or non-social features (arrows, words and objects). By contrast, the results are inconsistent with theories that suggest that the control of imitation relies on two nodes of the ToM network (medial prefrontal cortex and right temporoparietal junction), which support a process of self-other distinction (Brass & Heyes, 2005; Brass et al., 2009).

The lack of evidence for domain-specificity also has important implications for researchers interpreting findings from the SRC task. Experimentally manipulating social

variables or neurostimulation procedures that target "social networks" could be influencing domain-general orienting and selection processes, rather than an imitative component specifically. In other words, modification to a domain-general system could be the route by which social control is exerted in imitation, much like other domains of social cognition and perception that show interplay between domain-general and specific systems (Zaki et al., 2010; Baldauf & Desimone, 2014). Some prior models of imitation have included roles for domain-general processes in the control of imitation (Cross et al., 2013; Heyes, 2011; Gowen & Poliakoff, 2012; Darda et al., 2018), but much more attention has been placed on models that include "social control", which rests on domain-specific processes (Brass et al., 2009; Heyes & Brass, 2005; Wang & Hamilton, 2012). Importantly, there is no reason why the cognitive structure proposed here, which places much greater emphasis on interactions between domain-general and domain-specific systems, could not also operate during live social interactions, such as when social factors influence overt copying behaviour (Chartrand & Lakin, 2013). That is, social factors that modulate real-world copying behaviour could also operate through a domain-general action selection mechanism. Therefore, before a moderating factor can be concluded to be operating on a domain-specific (imitative) process, it would be important to clearly identify the component parts of the broader imitative mechanism as well as which parts are influenced by social factors. In addition, alternative neurocognitive models are required, which do not rely on right temporoparietal junction and medial prefrontal cortex as solitary imitation controllers (Brass et al., 2009; Wang & Hamilton, 2012), but instead consider interplay between domain-general and domain-specific systems. Such future directions for imitation research are the focus of the next section.

4. Looking forward

Theory and methods in imitation research need updating if the use of SRC paradigms are going to be able to contribute to a progressive research programme (Lakatos, 1970), and thus enable a cumulative and reproducible science to emerge (Munafo et al., 2017). In the following two subsections, I outline ideas for future research in terms of opportunities for theoretical and methodological development.

4.1 Validity and domain-specificity

The level of validity and domain-specificity of the component processes of SRC measures of automatic imitation need to be systematically investigated. Validity and specificity are orthogonal dimensions and could both range from low to high (Figure 3). Validity in this sense would be a composite of multiple dimensions of validity. Both validity and domain-specificity should be evaluated by degree: To what extent are component processes, such as action representation and control, valid indices of processes associated with copying 'in the wild'? And to what extent does automatic imitation recruit processes that rely on domain-specific, domain-general or a combination of cognitive architectures?

Four broad combinations of validity and specificity emerge. Figure 3 illustrates these dimensions as positions A to D. Position A represents high validity and domain-specificity. This would suggest that the task indexes a measure that is associated with imitation and relies on a cognitive and neural architecture that is relatively domain-specific, such that it preferentially operates in social contexts and when interacting with other people. This is similar to the position put forward by a number of researchers, at least implicitly (Brass et al., 2009; de Guzman et al., 2016; Sowden & Shah, 2014; Spengler et al., 2009). Position B represents low validity, but high domain-specificity. This would suggest that the process does not operate in overt copying behaviour, but it is specific to social interactions rather than a more general cognitive mechanism. This would suggest that the process is tied to the task and

socialness but is not engaged in overt copying behaviour. Position C represents high validity, but low specificity. This would suggest that the process is tied to overt copying behaviour, but that it is a domain general mechanism. An example could be a general selection mechanism that takes all kinds of different inputs and prioritises them in the context of an individual's current goal. Finally, position D would show low validity and low specificity. This would suggest that the process is not tied to imitation particularly, nor is it tied to a specific social domain of operation. An example could be a general selection mechanism that operates in the task but does not reflect the type of selection that operates in live copying behaviour. Position D would be worth confirming to suggest that other tasks may be more appropriately suited to study imitation.

As outlined in the introduction, automatic imitation is likely be a product of several component processes such as action representation and action control (Figure 2). Component processes such as action representation and control are not required to have the same profile of validity and specificity. For example, action representation could rely to a greater extent on domain-specific resources, whereas control could rely to a greater extent on domain-general resources. Furthermore, component processes need not operate the same in SRC and real-world measures of imitation. For example, it is plausible that SRC tasks measure action representation processes that are also used in real world copying behaviours, but domain-general selection processes that are not used in real-world copying. Alternatively, it is also plausible that the basic component processes operate in a largely similar manner in SRC and real-world contexts, but the interaction between these processes differs. Indeed, the type of control over action representations exerted during live social interactions is likely to be much richer and more diverse than the socially impoverished SRC context. As a consequence, real-world imitation could be underpinned by quantitatively stronger interactions between component processes (e.g., stronger executive influences on action representations in real-life

than SRC tasks) as well as qualitatively different kinds of interaction between component processes (e.g., general executive processes could be recruited more in the SRC task, but more socially-specific forms of control, which may result from outputs of theory of mind systems, may be recruited in real-world settings). Therefore, one key focus of future research should be to address where component processes lie in this validity-specificity space, rather than where tasks as a whole lie. Investigating how social context and the type of imitation (SRC vs. real-world) influences the operation of component processes would be a valuable contribution to the literature. In short, consistent with recent suggestions in social cognition (Schaafsma et al., 2015), it would be important to deconstruct automatic imitation into component pieces, study them and then reconstruct them into a better understanding of the overall, complex phenomenon of interest – automatic imitation. By doing so, we may develop a better idea of the structure of the processes involved. The next section focusses on testing more complex cognitive and brain models.

4.2 Cognitive and brain models need revision and higher complexity

In scientific research, it is important that theoretical foundations are clearly articulated, as these provide the blueprints for research programmes to generate and test hypotheses (Gray, 2017; Meehl, 1990; Popper, 1959). To develop more detailed cognitive and neurobiological models of imitation, several theoretical avenues are ripe for further investigation. For instance, considering the issue of domain-specificity, which was outlined in the previous section, it is noteworthy that one set of processes associated with imitation could be domain-general (control / selection) without requiring the entire set of processes that support imitation to be domain-general (Figure 4). For example, it is possible that we use a specialised neural architecture for detecting the presence of conspecifics and their actions (Kanwisher, 2010; Caspers et al., 2010), but select between different actions using domain-

general control systems (Duncan, 2010; Figure 4). This structure closely follows the structure put forward in the 'controlled semantic cognition' view of semantic cognition (Lambon Ralph et al., 2017). This type of analysis supports the proposal that information processing models of social cognition need not be exclusively tied to component processes that are domain-specific (Spunt & Adolphs, 2017). The analysis also highlights the importance of clearly defining component processes in a model, as well as the possible information processing specificity of each component. Relatedly, it may well be the links between domain-specific (person perception and action observation) and domain-general processes (selection) that make a process distinctly social rather than non-social. Indeed, positing links between domain-specific and general systems has proven useful in other domains of social perception and cognition (Zaki et al., 2010; Baldauf & Desimone, 2014; Quadflieg et al., 2010; Sreenivasan et al., 2014; Spunt & Adolphs, 2017).

Prior models of imitation have proposed links between domain-general and domain-specific systems (Cross et al., 2013; Heyes, 2011; Gowen & Poliakoff, 2012). Other models of imitation emphasise social control exerted by domain-specific systems such as the theory of mind network, but do not mention domain-general control architectures (Brass et al., 2009; Wang & Hamilton, 2012). However, none of these models make reference to representation and control, as defined in semantic cognition (Jefferies, 2013; Lambon Ralph et al., 2017). As such, I suggest that it is valuable to at least consider these different models of imitation with respect to the definitions of representation and control outlined in section 2.1 (Table 1). The breakdown of these models in terms of representation and control (as outlined in Table 1) is only suggestive at this stage and is likely to underestimate the complexity of mental processes involved. However, it is clear that these accounts make substantially different claims about the underlying neural bases of automatic imitation. As a consequence, Table 1 represents an

attempt to try to compare and contrast these accounts using a common language, in order to explore if these accounts generate distinct and testable hypotheses.

Although several recent models mention links between domain-general and domain-specific systems (Heyes, 2011; Gowen & Poliakoff, 2012), the model put forward in Figure 4, which is adapted from the 'controlled semantic cognition' framework (Lambon Ralph et al., 2017), differs from these models in at least two important aspects. First, recent models have a dual-route structure (Heyes, 2011; Gowen & Poliakoff, 2012), which separate a fast and efficient route from a slower and more controlled route. Although dual-route structures have gained prominence in social cognition recently, such as in the domain of theory of mind (Apperly & Butterfill, 2009), the same data can be explained with a single route structure (Carruthers, 2016; 2017). Therefore, the model proposed in Figure 4 has a single route structure, which always has a degree of representation and control. The amount and type of information represented and the degree of control exerted will vary based on context. In some contexts, there will be minimal control, whereas in other contexts, there will be greater control exerted. But in all cases, as outlined in semantic cognition (Lambon Ralph et al., 2017), imitative control will involve the same interacting systems of representation and control.

A second distinction of the model proposed in Figure 4 compared to prior models of imitative processes is the emphasis on distributed processing. Some prior models of imitation do not cover the likely range of mechanisms and brain circuits involved (see Table 1). A consequence of having a restricted set of components in a model of brain function, is the potential to over-ascribe cognitive function to segregated and localised processes. For example, Hauser & Hickok (2010; Hickok, 2013) propose that a key functional property of the mirror neuron system is a role in action selection without mentioning complementary roles for other social or non-social brain circuits. It is reasonable to propose a role for the

mirror neuron system in action selection, as well as other processes, but it seems unlikely that it operates alone. Furthermore, models that do not include domain-general control, may be biased towards an over-emphasis on domain-specific control (Brass et al., 2009; Wang & Hamilton, 2012). Indeed, findings taken to support a domain-specific "social control" view of the theory of mind network in imitation (Wang, Ramsey & Hamilton, 2011; Wang & Hamilton, 2012), may reflect an updated person representation, rather than a form of control. Future research should test between these possibilities by including a domain-general control component in future models. In short, something as complex as action selection in imitative contexts is likely to require a distributing and interacting set of cognitive and neural processes spanning domain-general and domain-specific processes.

Even in distributed models, such as Gowen & Poliakoff's (2012) dual route structure, which spans several brain circuits, a cognitively elaborate set of processes is still ascribed to localised structures. For example, Gowen & Poliakoff's (2012) proposed role for the mirror neuron system is "The visuomotor route consists of parietal and frontal circuits involving the translation of visual information into action via SRC mechanisms which produces visuomotor priming by facilitating or interfering with the programming of the participant's response." This may be true, but it also need not be the case. Cognitively elaborate computations, such as those ascribed to the mirror neuron system, could remain distributed and taken on by a set of interacting brain circuits and systems. For instance, the model of imitative control outlined in Figure 4 is always a product of representation and control, but the content and relative weights assigned to each component can vary. Of course, future research is required to test between these possibilities. More generally, emphasising the integration of information across distributed architectures (Bullmore & Sporns, 2009), may have a positive impact on continuing debates regarding functional claims over segregated circuits, such as the mirror neuron system (Hickok, 2009; Rizzolatti & Sinigaglia, 2010; 2016).

The model proposed in Figure 4 is labelled the 'controlled social cognition' account of automatic imitation due to its direct adaption from the 'controlled semantic cognition' framework (Jefferies, 2013; Lambon Ralph et al., 2017). As Table 1 outlines, alternatives to this structure exist and offer fertile ground for future studies to distinguish between these accounts. However, one advantage of the model put forward here is that its core features are supported by a wealth of evidence in the domain of semantic cognition, which spans neuropsychology, fMRI and neurostimulation techniques. This evidence shows clear distinctions between representation and control, in terms of underlying neural structures (for reviews, see Jefferies, 2013; Lambon Ralph et al., 2017). The model presented here, therefore, proposes that sociocognitive processes, such as those involved in imitative contexts, rely on the same broad semantic system that underpins other complex cognitive processes. Furthermore, the structure of that semantic system has two principal and interacting components of representation and control, which rely on largely distinct brain circuits.

An alternative approach to theory development can be provided by considering specialised recruitment of domain-general architectures. For instance, a recent fMRI study showed that a region within dorsomedial prefrontal cortex, which shows a domain-general response, also shows a response that is tuned for making specifically social attributions (Spunt & Adolphs, 2015). Spunt and Adolphs (2015) interpret this result as the specialised or "expert" use for social purposes of a largely domain-general system. Finally, a greater focus on computational modelling approaches to automatic imitation (Cooper et al., 2013), as well as theories that posit links between biological and functional levels of description (Cook et al., 2014), could enable more precise quantitative predictions to be made thus providing a stricter test of underlying hypotheses (Meehl, 1990). Taken together, these theoretical considerations, as well as recent empirical findings, suggest that greater complexity is required when building theories of social cognition that extend beyond a strict division

between general and specific systems (Spunt & Adolphs, 2017). Moreover, online tools are available, such as Theory Mapping (Gray, 2017; www.theorymaps.org), which aid in the development and comparison of theories by enabling researchers to visualise theories and model spaces in a common language and store them online (Supplementary Figure 1).

The proposed avenues for theory development are not only applicable to imitation research, but instead generalise to other domains of social cognition. Indeed, similar debates regarding domain-specificity have arisen in gaze perception and perspective taking literatures that also rely on SRC paradigms (Frischen et al., 2007; Santiesteban et al., 2014). As a consequence, core organising principles from the model outlined for imitation in Figure 4, whereby the input is social, but the selection mechanism may be domain-general, could be applicable in these contexts also. For instance, in order to perceive eye gaze or take another's perspective, it is essential to recognise the presence of another agent, but nonetheless, conflict resolution between gaze cues or perspectives may be resolved through a domain-general selection mechanism. Moreover, it would be the link between perceptual and selection processes that distinguishes social cognition from non-social cognition. In other words, just as has been outlined in memory research (Cabeza & Moscovitch, 2013), specialisation for social cognition may be characterised by functional circuits that consist of multiple components, without the need for each component to be domain-specific (Spunt & Adolphs, 2017).

4.3 More robust methods – power, replication, meta-analysis, pre-registration and open data

In addition to establishing reliability and validity of SRC imitation paradigms, this field would also benefit substantially from integrating practices emerging from the recent movement toward open science more generally. I do not provide a detailed account of the motivation and means to improve rigour and reproducibility, as these issues have been

covered at length by others (Cumming, 2012; Simmons et al., 2011; 2018; Nelson et al., 2018; Munafo et al., 2017; Zwaan et al., 2017). Instead, I wish to highlight that imitation research, like the fields of psychology and neuroscience more generally (Button et al., 2013; Open Science Collaboration, 2015), could benefit from revisiting foundational methodological approaches. A cycle of studies with low power and a publication bias skewed towards positive results has produced a questionable level of evidence for many social claims made regarding imitation. Approaches that include pre-registering studies, appropriately powered designs, replicating results, using meta-analyses and making datasets open and available to others would be of great benefit to imitation research. Some of these approaches are already being taken in SRC research on automatic imitation (Butler et al., 2015; 2016; Cracco, Bardi et al., 2018; Cracco, Genschow, et al., 2018; Darda et al., 2018; Genschow et al., 2017), but more universal adoption of these methodological principles would provide considerable benefits to current and future researchers working in this area. These methodological steps go hand-in-hand with improved theoretical models and are a necessary starting point to improve the evidence base for understanding the mechanisms associated with automatic imitation.

Conclusion

It currently remains unclear what, precisely, reaction time indices of automatic imitation measure, due to a lack of research that directly assesses dimensions of validity and domain-specificity. It terms of validity, there is insufficient empirical evidence that demonstrates the extent to which reaction time measures index processes that are associated with overt copying behaviours. The lack of validity means that it is possible that reaction time measures are indexing a different aspect of imitation that is not associated with overt copying behaviour or something entirely unrelated to imitation. Furthermore, it terms of domain-specificity, much

like other complex cognitive and social processes, reaction time measures are likely to reflect a combination of component processes. Such component processes could operate on mechanisms that are specifically tied to social cognition and imitation, more general mechanisms or a combination of both types of mechanism.

Based on this analysis of the current literature, I suggest that researchers should be more cautious when drawing parallels between reaction time measures of automatic imitation and overt copying behaviours. In addition, three suggestions are made that aim to develop understanding of automatic imitation using SRC paradigms. First, studies that assess the validity of SRC paradigms as an index of automatic imitation are required. Second, the extent that component processes of automatic imitation rely on domain-specific, domain-general or a combination of both types of cognitive system needs establishing with greater precision. Moreover, theories of imitation that move beyond neat divisions between domain-specific and domain-general systems should be given greater emphasis. Third, a greater focus on methodological rigour is required to enable a cumulative science to develop. In sum, the SRC index of automatic imitation holds much promise as a tool to study component processes associated with automatic imitation and thus answer questions at the intersection of psychology and cognitive neuroscience. By providing a critical review of the current state of the art, this paper aims to open a discussion regarding optimal methods to study automatic imitation using reaction time measures, which will hopefully lead to new research questions, theories and methodological approaches.

Acknowledgements

I would like to thank Emily Cross, Kohinoor Darda, Paul Downing and four anonymous reviewers for comments on earlier versions of this manuscript.

References

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states?. Psychological review, 116(4), 953.

Baldauf, D., & Desimone, R. (2014). Neural Mechanisms of Object-Based Attention. Science, 344(6182), 424-427. doi: 10.1126/science.1247003

Barrett, H. C. (2012). A hierarchical model of the evolution of human brain specializations. *Proceedings of the National Academy of Sciences, 109*(Supplement 1), 10733-10740. doi: 10.1073/pnas.1201898109

Bertenthal, B. I., Longo, M. R., & Kosobud, A. (2006). Imitative response tendencies following observation of intransitive actions. Journal of Experimental Psychology: Human Perception and Performance, 32(2), 210-225. doi: 10.1037/0096-1523.32.2.210

Bien, N., Roebroeck, A., Goebel, R., & Sack, A. T. (2009). The Brain's Intention to Imitate: The Neurobiology of Intentional versus Automatic Imitation. Cerebral Cortex, 19(10), 2338-2351. doi: 10.1093/cercor/bhn251

Boyer, T. W., Longo, M. R., & Bertenthal, B. I. (2012). Is automatic imitation a specialized form of stimulus,Äìresponse compatibility? Dissociating imitative and spatial compatibilities. Acta Psychologica, 139(3), 440-448. doi: 10.1016/j.actpsy.2012.01.003

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychol (Amst), 106*(1-2), 3-22.

Brass, M., Bekkering, H., Wohlschlager, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: comparing symbolic, spatial, and imitative cues. Brain Cogn, 44(2), 124-143.

Brass, M., Derrfuss, J., & von Cramon, D. Y. (2005). The inhibition of imitative and overlearned responses: a functional double dissociation. Neuropsychologia, 43(1), 89-98.

Brass, M., & Heyes, C. (2005). Imitation: is cognitive neuroscience solving the correspondence problem? Trends Cogn Sci, 9(10), 489-495.

Brass, M., Ruby, P., & Spengler, S. (2009). Inhibition of imitative behaviour and social cognition. Philosophical Transactions of the Royal Society B: Biological Sciences, 364(1528), 2359-2367. doi: 10.1098/rstb.2009.0066

Brass, M., Zysset, S., & von Cramon, D. Y. (2001). The inhibition of imitative response tendencies. NeuroImage, 14(6), 1416-1423. doi: 10.1006/nimg.2001.0944

Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, *10*(3), 186.

Butler, E. E., Ward, R., & Ramsey, R. (2015). Investigating the Relationship between Stable Personality Characteristics and Automatic Imitation. PLoS ONE, 10(6). doi: ARTN e012965110.1371/journal.pone.0129651

Butler, E. E., Ward, R., & Ramsey, R. (2016). The Influence of Facial Signals on the Automatic Imitation of Hand Actions. *Frontiers in Psychology, 7*(1653). doi: 10.3389/fpsyg.2016.01653

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafo, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci, 14(5), 365-376. doi: 10.1038/nrn3475

Cabeza, R., & Moscovitch, M. (2013). Memory Systems, Processing Modes, and Components: Functional Neuroimaging Evidence. Perspectives on Psychological Science, 8(1), 49-55. doi: 10.1177/1745691612469033

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56(2), 81-105.
Carruthers, P. (2016). Two systems for mindreading?. *Review of Philosophy and Psychology*, *7*(1), 141-162.
Carruthers, P. (2017). Mindreading in adults: evaluating two-systems views. *Synthese*, *194*(3), 673-688.
Caspers, S., Zilles, K., Laird, A. R., & Eickhoff, S. B. (2010). ALE meta-analysis of action observation and imitation in the human brain. NeuroImage, 50(3), 1148-1167.
Catmur, C., & Heyes, C. (2011). Time course analyses confirm independence of imitative and spatial compatibility. Journal of Experimental Psychology: Human Perception and Performance, 37(2), 409-421. doi: 10.1037/a0019325
Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: the perception-behavior link and social interaction. J Pers Soc Psychol, 76(6), 893-910.
Chartrand, T. L., & Lakin, J. L. (2013). The Antecedents and Consequences of Human Behavioral Mimicry. Annual Review of Psychology, 64(1), 285-308. doi: doi:10.1146/annurev-psych-113011-143754
Cook, J., & Bird, G. (2011). Social attitudes differentially modulate imitation in adolescents and adults. *Experimental Brain Research, 211*(3), 601-612. doi: 10.1007/s00221-011-2584-4
Cook, R., Bird, G., Catmur, C., Press, C., & Heyes, C. (2014). Mirror neurons: from origin to function. *Behav Brain Sci, 37*(2), 177-192. doi: 10.1017/S0140525X13000903
Cooper, R. P., Catmur, C., & Heyes, C. (2013). Are automatic imitation and spatial compatibility mediated by different processes? *Cogn Sci, 37*(4), 605-630. doi: 10.1111/j.1551-6709.2012.01252.x
Corbetta, M., Patel, G., & Shulman, G. L. (2008). The Reorienting System of the Human Brain: From Environment to Theory of Mind. Neuron, 58(3), 306-324.
Cracco, E., Bardi, L., Desmet, C., Genschow, O., Rigoni, D., Coster, L. D., . . . Brass, M. (2018). Automatic Imitation: A Meta-Analysis. *Psychological Bulletin, 144*(5), 453-500.
Cracco, E., Genschow, O., Radkova, I., & Brass, M. (2018). Automatic imitation of pro-and antisocial gestures: Is implicit social behavior censored?. *Cognition*, *170*, 179-189.
Cronbach, L. J. (1990). Essentials of psychological testing (5th ed.). New York: Harper & Row.
Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. Psychological Bulletin, 52(4), 281-302.
Cross, K. A., & Iacoboni, M. (2013). Optimised neural coding? Control mechanisms in large cortical networks implemented by connectivity changes. Human Brain Mapping, 34, 213–225.
Cross, K. A., Torrisi, S., Losin, E. A. R., & Iacoboni, M. (2013). Controlling automatic imitative tendencies: Interactions between mirror neuron and cognitive control systems. Neuroimage, 83,493–504.
Cumming, G. (2012). Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis. New York: Routledge.
Darda, K. M., Butler, E. E., & Ramsey, R. (2018). Functional Specificity and Sex Differences in the Neural Circuits Supporting the Inhibition of Automatic Imitation. Retrieved from psyarxiv.com/rvh56.
de Guzman, M., Bird, G., Banissy, M. J., & Catmur, C. (2016). Self–other control processes in social cognition: from imitation to empathy. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*(1686). doi: 10.1098/rstb.2015.0079

Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends Cog Sci, 14*(4), 172-179. doi: 10.1016/j.tics.2010.01.004
Farrer, C., & Frith, C. D. (2002). Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency. *NeuroImage, 15*(3), 596-603.
Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. Proceedings of the National Academy of Sciences, U.S.A., 110, 16616–16621.
Fodor, J. A. (1983). *The Modularity of Mind*. Cambridge, MA: The MIT Press.
Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. Psychological Bulletin, 133(4), 694-724. doi: 10.1037/0033-2909.133.4.694
Frith, C. D., & Frith, U. (1999). Interacting minds--a biological basis. Science, 286(5445), 1692-1695.
Genschow, O., van Den Bossche, S., Cracco, E., Bardi, L., Rigoni, D., & Brass, M. (2017). Mimicry and automatic imitation are not correlated. PLoS ONE, 12(9), e0183784. doi: 10.1371/journal.pone.0183784
Gowen, E., Bolton, E., & Poliakoff, E. (2016). Believe it or not: Moving non-biological stimuli believed to have human origin can be represented as human movement. Cognition, 146, 431-438. doi: 10.1016/j.cognition.2015.10.010
Gowen, E., & Poliakoff, E. (2012). How does visuomotor priming differ for biological and non-biological stimuli? A review of the evidence. *Psychological research*, *76*(4), 407-420.
Gray, K. (2017). How to Map Theory: Reliable Methods Are Fruitless Without Rigorous Theory. Perspect Psychol Sci, 12(5), 731-741. doi: 10.1177/1745691617691949
Gleibs, I. H., Wilson, N., Reddy, G., & Catmur, C. (2016). Group Dynamics in Automatic Imitation. *PLoS ONE, 11*(9), e0162880. doi: 10.1371/journal.pone.0162880
Heyes, C. (2011). Automatic imitation. Psychological Bulletin, 137(3), 463-483. doi: 10.1037/a0022288
Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of cognitive neuroscience*, *21*(7), 1229-1243.
Hickok, G. (2013). Do mirror neurons subserve action understanding? *Neuroscience Letters, 540*, 56-58. doi: https://doi.org/10.1016/j.neulet.2012.11.001
Hickok, G., & Hauser, M. (2010). (Mis)understanding mirror neurons. *Current Biology, 20*(14), R593-R594. doi: https://doi.org/10.1016/j.cub.2010.05.047
Hogeveen, J., & Obhi, S. S. (2013). Automatic imitation is automatic, but less so for narcissists. Experimental Brain Research, 224(4), 613-621. doi: 10.1007/s00221-012-3339-6
Iacoboni, M. (2009). Imitation, Empathy, and Mirror Neurons. *Annual Review of Psychology, 60*(1), 653-670. doi: doi:10.1146/annurev.psych.60.110707.163604
Jefferies, E. (2013). The neural basis of semantic cognition: converging evidence from neuropsychology, neuroimaging and TMS. Cortex, 49(3), 611-625.
Kanwisher, N. (2010). Functional specificity in the human brain: A window into the functional architecture of the mind. *Proceedings of the National Academy of Sciences, 107*(25), 11163-11170. doi: 10.1073/pnas.1005062107
Kilner, J. M., Paulignan, Y., & Blakemore, S. J. (2003). An interference effect of observed biological movement on action. *Curr Biol, 13*(6), 522-525.
Klapper, A., Ramsey, R., Wigboldus, D., & Cross, E. S. (2014). The Control of Automatic Imitation Based on Bottom–Up and Top–Down Cues to Animacy: Insights from Brain

and Behavior. *Journal of Cognitive Neuroscience, 26*(11), 2503-2513. doi: 10.1162/jocn_a_00651
Lakatos, I. (1970). Falsification and the methodology of scientific research programme. In I. Lakatos & A. Musgrave (Eds.), Criticism and the growth of knowledge (pp. 91–196). Cambridge, UK: Cambridge University Press.
Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. Nature Reviews Neuroscience, 18(1), 42.
Leighton, J., Bird, G., Orsini, C., & Heyes, C. (2010). Social attitudes modulate automatic imitation. *Journal of Experimental Social Psychology, 46*(6), 905-910. doi: https://doi.org/10.1016/j.jesp.2010.07.001
Likowski, K. U., Weyers, P., Seibt, B., Stöhr, C., Pauli, P., & Mühlberger, A. (2011). Sad and Lonely? Sad Mood Suppresses Facial Mimicry. Journal of Nonverbal Behavior, 35(2), 101-117. doi: 10.1007/s10919-011-0107-4
Marsh, L. E., Bird, G., & Catmur, C. (2016). The imitation game: Effects of social cues on 'imitation' are domain-general in nature. NeuroImage, 139, 368-375. doi: 10.1016/j.neuroimage.2016.06.050
McIntosh, D. N. (2006). Spontaneous facial mimicry, liking and emotional contagion. Polish Psychological Bulletin, 37(1), 31-42.
Meehl, P. E. (1990). Appraising and Amending Theories: The Strategy of Lakatosian Defense and Two Principles that Warrant It. Psychological Inquiry, 1(2), 108-141. doi: 10.1207/s15327965pli0102_1
Michael, J., & D'Ausilio, A. (2015). Domain-specific and domain-general processes in social perception – A complementary approach. *Consciousness and Cognition, 36*, 434-437. doi: http://dx.doi.org/10.1016/j.concog.2014.12.009
Molenberghs, P., Cunnington, R., & Mattingley, J. B. (2012). Brain regions with mirror properties: A meta-analysis of 125 human fMRI studies. *Neuroscience & Biobehavioral Reviews, 36*(1), 341-349. doi: 10.1016/j.neubiorev.2011.07.004
Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., . . . Ioannidis, J. P. A. (2017). A manifesto for reproducible science. Nature Human Behaviour, 1, 0021. doi: 10.1038/s41562-016-0021
Nelson, L. D., Simmons, J., & Simonsohn, U. (2018). Psychology's Renaissance. Annu Rev Psychol, 69, 511-534. doi: 10.1146/annurev-psych-122216-011836
Obhi, S. S., Hogeveen, J., Giacomin, M., & Jordan, C. H. (2014). Automatic imitation is reduced in narcissists. Journal of Experimental Psychology: Human Perception and Performance, 40(3), 920-928. doi: 10.1037/a0034056
Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251). doi: 10.1126/science.aac4716
Pan, X., & Hamilton, A. F. de C. (2018). Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. British Journal of Psychology, 109(3), 395-417. doi: doi:10.1111/bjop.12290
Poldrack, R., Kittur, A., Kalar, D., Miller, E., Seppa, C., Gil, Y., . . . Bilder, R. (2011). The Cognitive Atlas: Toward a Knowledge Foundation for Cognitive Neuroscience. Frontiers in Neuroinformatics, 5(17). doi: 10.3389/fninf.2011.00017
Popper, K. R. (1959). *The Logic of Scientific Discovery*. London: Hutchinson.
Press, C. (2011). Action observation and robotic agents: Learning and anthropomorphism. *Neuroscience & Biobehavioral Reviews, 35*(6), 1410-1418. doi: 10.1016/j.neubiorev.2011.03.004
Press, C., Bird, G., Flach, R., & Heyes, C. (2005). Robotic movement elicits automatic imitation. *Brain Res Cogn Brain Res, 25*(3), 632-640.

Quadflieg, S., Flannigan, N., Waiter, G. D., Rossion, B., Wig, G. S., Turk, D. J., & Macrae, C. N. (2011). Stereotype-based modulation of person perception. NeuroImage, 57(2), 549-557. doi: 10.1016/j.neuroimage.2011.05.004

Ray, E., & Heyes, C. (2011). Imitation in infancy: the wealth of the stimulus. Developmental Science, 14(1), 92-105. doi: 10.1111/j.1467-7687.2010.00961.x

Rizzolatti, G., & Craighero, L. (2004). The Mirror-Neuron System. *Annu Rev Neurosci, 27*, 169-192.

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature reviews neuroscience*, *11*(4), 264.

Rizzolatti, G., & Sinigaglia, C. (2016). The mirror mechanism: a basic principle of brain function. *Nature Reviews Neuroscience*, *17*(12), 757.

Ruby, P., & Decety, J. (2001). Effect of subjective perspective taking during simulation of action: a PET investigation of agency. *Nat Neurosci, 4*(5), 546-550.

Russell, J. A. (1978). Evidence of convergent validity on the dimensions of affect. Journal of Personality and Social Psychology, 36(10), 1152-1168. doi: 10.1037/0022-3514.36.10.1152

Santiesteban, I., Catmur, C., Hopkins, S. C., Bird, G., & Heyes, C. (2014). Avatars and arrows: Implicit mentalizing or domain-general processing? Journal of Experimental Psychology: Human Perception and Performance, 40(3), 929-937. doi: 10.1037/a0035175

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". NeuroImage, 19(4), 1835-1842.

Schaafsma SM, Pfaff DW, Spunt RP, & Adolphs R (2015). Deconstructing and Reconstructing “Theory of Mind”. *Trends in Cognitive Sciences*. 19(2), 65-72.

Schuwerk, T., Schurz, M., Muller, F., Rupprecht, R., & Sommer, M. (2017). The rTPJ's overarching cognitive function in networks for attention and theory of mind. *Soc Cogn Affect Neurosci, 12*(1), 157-168. doi: 10.1093/scan/nsw163

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. Psychological Science, 22(11), 1359-1366. doi: 10.1177/0956797611417632

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2018). False-positive citations. *Perspectives on Psychological Science*. *13*(2), 255-259.

Simon, J. R. (1969). Reactions toward the source of stimulation. *J Exp Psychol, 81*(1), 174-176.

Sowden, S., Brewer, R., Catmur, C., & Bird, G. (2016). The specificity of the link between alexithymia, interoception, and imitation. Journal of Experimental Psychology: Human Perception and Performance, 42(11), 1687-1692. doi: 10.1037/xhp0000310

Sowden, S., & Shah, P. (2014). Self-other control: a candidate mechanism for social cognitive function. *Frontiers in Human Neuroscience, 8*(789). doi: 10.3389/fnhum.2014.00789

Spengler, S., von Cramon, D. Y., & Brass, M. (2009). Control of shared representations relies on key processes involved in mental state attribution. *Human Brain Mapping, 30*(11), 3704-3718. doi: 10.1002/hbm.20800

Spunt, R. P., & Adolphs, R. (2015). Folk Explanations of Behavior: A Specialized Use of a Domain-General Mechanism. Psychological Science, 26(6), 724-736. doi: 10.1177/0956797615569002

Spunt, R. P., & Adolphs, R. (2017). A new look at domain specificity: insights from social neuroscience. *Nat Rev Neurosci, 18*(9), 559-567. doi: 10.1038/nrn.2017.76

Sreenivasan, K. K., Gratton, C., Vytlacil, J., & D'Esposito, M. (2014). Evidence for working memory storage operations in perceptual cortex. Cogn Affect Behav Neurosci, 14(1), 117-128. doi: 10.3758/s13415-013-0246-7

Stel, M., van Baaren, R. B., Blascovich, J., van Dijk, E., McCall, C., Pollmann, M. M. H., . . . Vonk, R. (2010). Effects of a priori liking on the elicitation of mimicry. Experimental Psychology, 57(6), 412-418. doi: 10.1027/1618-3169/a000050

Stürmer, B., Aschersleben, G., & Prinz, W. (2000). Correspondence effects with manual gestures and postures: A study of imitation. *Journal of Experimental Psychology: Human Perception and Performance, 26*(6), 1746-1759.

van Baaren, R., Janssen, L., Chartrand, T. L., & Dijksterhuis, A. (2009). Where is the love? The social aspects of mimicry. Philosophical Transactions of the Royal Society B: Biological Sciences, 364(1528), 2381-2389. doi: 10.1098/rstb.2009.0057

van Baaren, R. B., Fockenberg, D. A., Holland, R. W., Janssen, L., & van Knippenberg, A. (2006). The Moody Chameleon: The Effect of Mood on Non–Conscious Mimicry. Social Cognition, 24(4), 426-437. doi: 10.1521/soco.2006.24.4.426

Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. Hum. Brain Mapp., 30, 829.

Wang, Y., & Hamilton, A. F. d. C. (2012). Social Top-down Response Modulation (STORM): A model of the control of mimicry in social interaction. *Frontiers in Human Neuroscience, 6*. doi: 10.3389/fnhum.2012.00153

Wang, Y., Ramsey, R., & de C. Hamilton, A. F. (2011). The Control of Mimicry by Eye Contact Is Mediated by Medial Prefrontal Cortex. *The Journal of Neuroscience, 31*(33), 12001-12010. doi: 10.1523/jneurosci.0845-11.2011

Yarkoni, T., & Westfall, J. (2017). Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. *Perspect Psychol Sci, 12*(6), 1100-1122. doi: 10.1177/1745691617693393

Zaki, J., Hennigan, K., Weber, J., & Ochsner, K. N. (2010). Social cognitive conflict resolution: contributions of domain-general and domain-specific neural systems. J Neurosci, 30(25), 8481-8488. doi: 10.1523/JNEUROSCI.0382-10.2010

Zwaan, R. A., Etz, A., Lucas, R. E., & Donnellan, M. B. (2017). Making Replication Mainstream. Behav Brain Sci, 1-50. doi: 10.1017/S0140525X17001972

Table 1. Representation (R) and control (C) across different accounts of automatic imitation.

Proponents	Cognitive and brain systems			
	Person perception network	**Mirror neuron system**	**Theory of Mind network**	**Multiple Demand network**
'Controlled social cognition' view (Figure 4; adapted from Lambon Ralph et al., 2017).	R	R	R	C
Brass et al., 2009; Heyes & Brass, 2005	R	R	C	?
Cross et al., 2013; Cross & Iacoboni, 2013	R	R	?	C
Gowen & Poliakoff, 2012	R	R & C	C	C
Hauser & Hickok, 2010; Hickok, 2013.	R	R & C	?	?
Heyes, 2011	R	R & C	C	C
Wang & Hamilton, 2012	R	R	C	?

Table 1. Representation (R) and control (C) across different accounts of automatic imitation.
The designation of representation and control for each account of automatic imitation is based on comparing the cited literature to the definitions of representation and control given in section 2.1, which follow the terms used in the domain of semantic cognition (Jefferies, 2013; Lambon Ralph et al., 2017). Given that the majority of authors cited above do not explicitly frame their accounts of imitation in terms of representation and control, the above positions are only suggestive and it is likely that each account may be more complex than this table is able to summarise. Nonetheless, it is certainly the case that different accounts either: A) make different claims about the roles of the above brain networks during automatic imitation; B) place considerably more or less emphasis on which brain networks are involved in controlling imitation, or; C) do not include certain brain networks in the control of imitation. As such, this table represents an attempt to try to compare and contrast these accounts using a common language, in order to find out if these accounts generate distinct and testable hypotheses. See section 2.1 for definitions of representation and control and a more detailed description of each brain network. For completeness and to aid comparison between accounts, I identify cases where roles for brain circuits were not specified using a question mark.
Abbreviations: R = representation; C = control; ? = unspecified.

Figure 1. Two approaches to measuring automatic imitation.

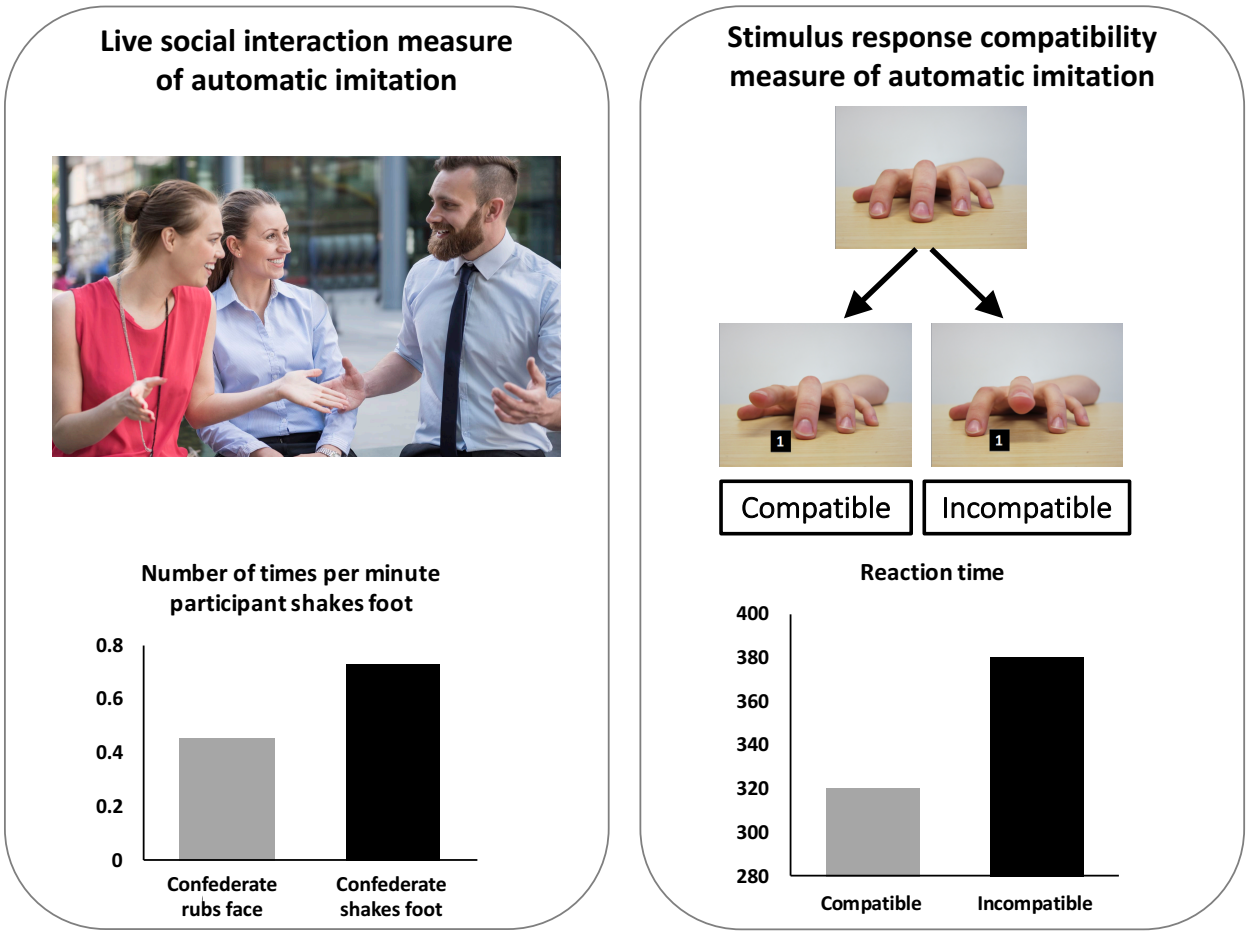


Figure 1. Two approaches to measuring automatic imitation are illustrated. In the left panel, unintentional copying behaviours are recorded during live social interactions (e.g., Chartrand & Bargh, 1999). A confederate may rub her face or shake her foot during a conversation, and the number of matching actions performed by the participant are recorded. The number of times per minute that a participant shakes her foot is greater when the confederate shakes her foot than rubs her face. As such, even though there is no instruction to copy actions, participants appear to unintentionally imitate the actions of their interaction partners. In the right panel, reaction times are recorded during a stimulus response compatibility paradigm (e.g., Brass et al., 2000). In one example paradigm, participants are instructed to lift their index finger if a number 1 is presented and lift their middle finger if a number 2 is presented. Concurrent with this instruction, participants also observe a finger lift that can either be compatible or incompatible with the participant's finger lift. Reaction times are longer in the incompatible than compatible conditions and this difference is thought to index the additional cognitive resources required to suppress the urge to imitate the incompatible action and enforce one's own intended action.

Figure 2. Brain networks associated with automatic imitation.

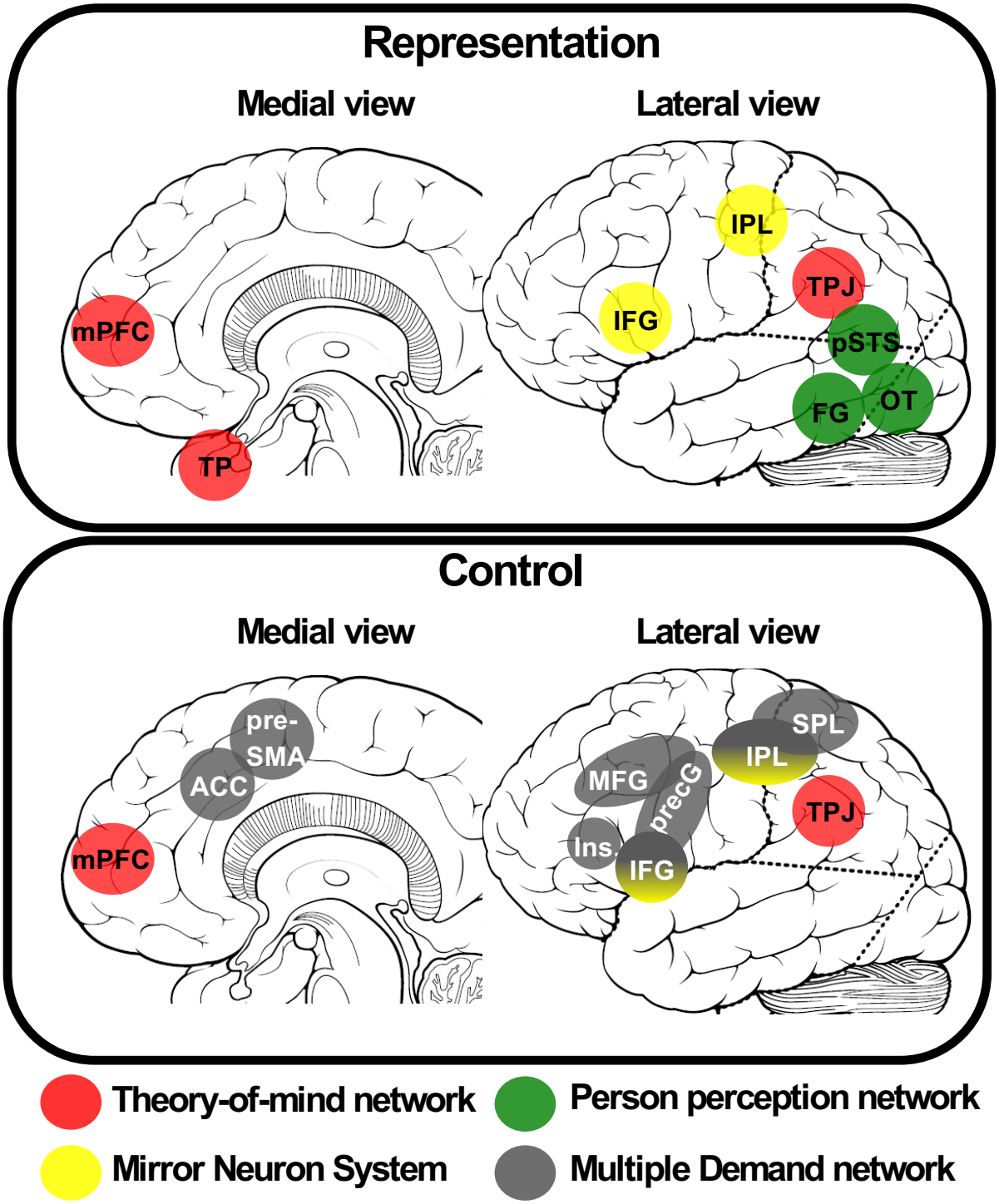


Figure 2. Brain networks associated with automatic imitation. The terms representation and control are used as defined in prior literature (see Section 2.1; Jefferies, 2013; Lambon Ralph et al., 2017). The representational system refers to knowledge of concepts, which are a product of higher-order relationships between sensory, motor, linguistic and affective sources of input. The control system manipulates the representational system as a function of the relevant context. In the context of automatic imitation, the representational content spans person and action representations including face, body, biological motion and action perception, as well as mental state attributions. As such, the neural substrates supporting such representations span fusiform and occipitotemporal cortices (Kanwisher, 2010), as well as the mirror neuron system (Caspers et al., 2010) and theory of mind network (Van Overwalle,

2009). In terms of control systems in automatic imitation, different accounts have been put forward, which differently emphasise roles for the multiple demand network (Cross et al., 2013; Darda et al., 2018), the mirror neuron system (Hauser & Kickok, 2010; Hickok, 2013) and the theory of mind network (Brass et al., 2009; Wang & Hamilton, 2012). Abbreviations: mPFC = medial prefrontal cortex; TP = temporal pole; IFG = inferior frontal gyrus; IPL = inferior parietal lobule; TPJ = temporoparietal junction; pSTS = posterior superior temporal sulcus; FG = fusiform gyrus; OT = occipitotemporal cortex; pre-SMA = pre-supplementary motor area; MFG = middle frontal gyrus; precG = precentral gyrus; SPL = superior parietal lobule; Ins. = insula; ACC = anterior cingulate cortex.

Figure 3. Structure of processing components in terms of validity and domain-specificity.

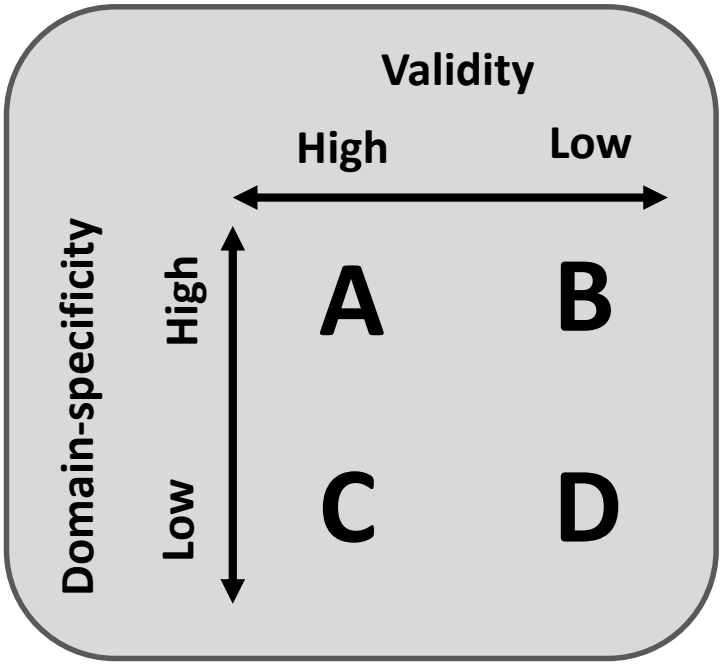


Figure 4. A 'controlled social cognition' account of automatic imitation.

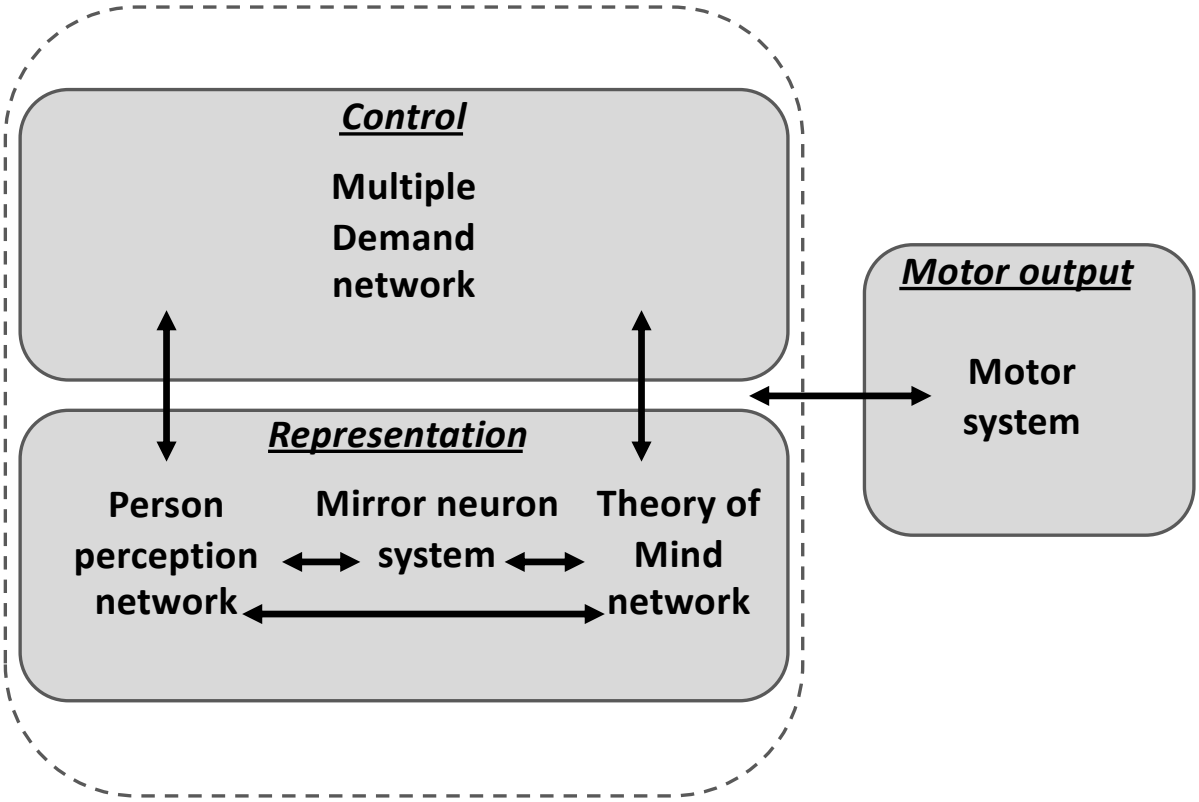
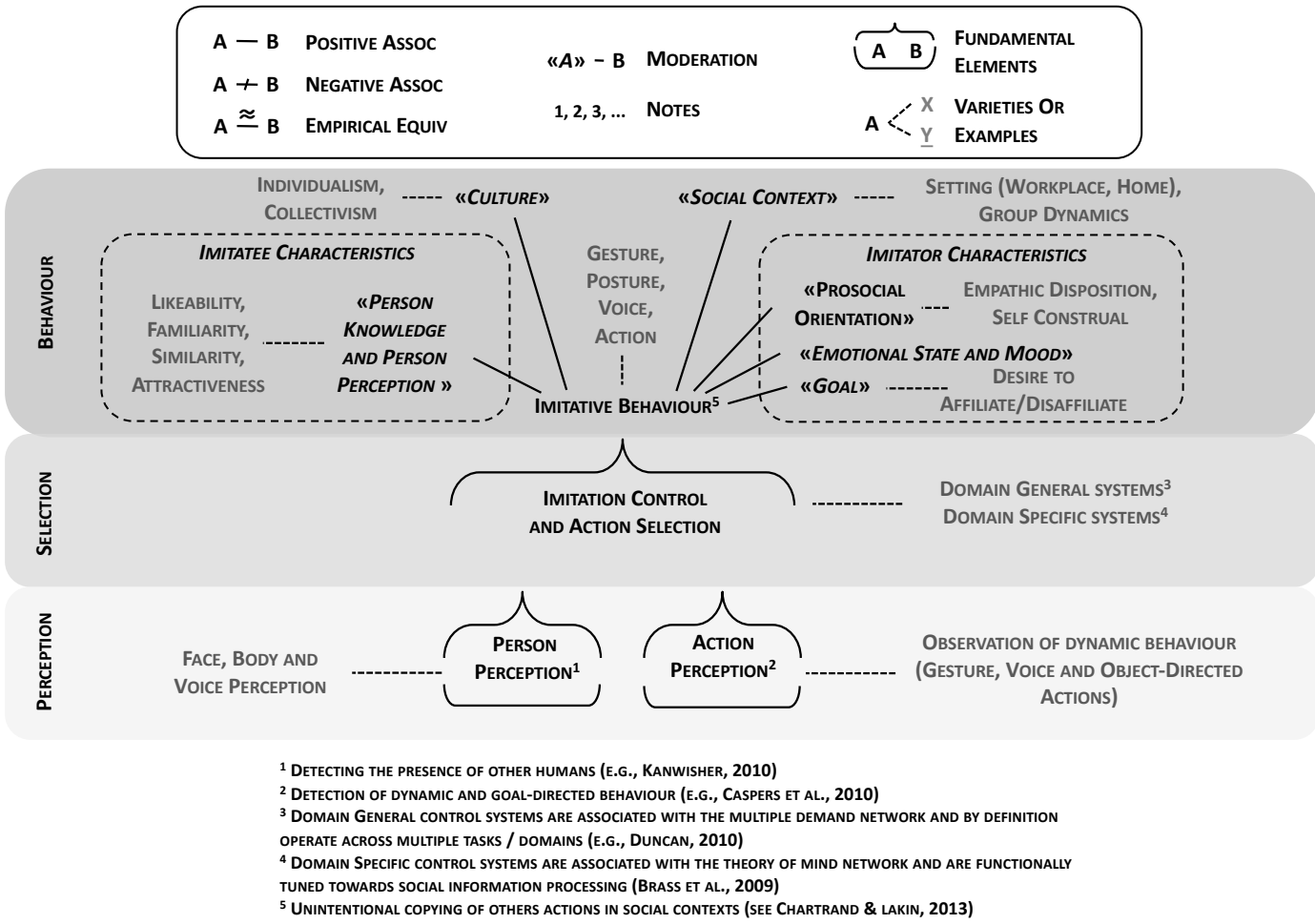


Figure 4. A 'controlled social cognition' account of automatic imitation. The proposed account of the cognitive and neural processes underpinning automatic imitation presented here is adapted from 'controlled semantic cognition' accounts of semantic cognition (Jefferies, 2013; Lambon Ralph et al., 2017). As Table 1 outlines, alternatives to this structure exist and offer fertile ground for future studies to distinguish between these accounts. However, one advantage of the model put forward here is that its core features are supported by a wealth of evidence in the domain of semantic cognition, which spans neuropsychology, fMRI and neurostimulation techniques. This evidence shows clear distinctions between representation and control, in terms of underlying neural structures, as visualised above (for reviews, see Jefferies, 2013; Lambon Ralph et al., 2017). The model presented here, therefore, proposes that sociocognitive processes, such as those involved in imitative contexts, rely on the same broad semantic system that underpins other complex cognitive processes. Furthermore, the structure of that semantic system has two principal and interacting components of representation and control, which rely on largely distinct brain circuits.

Supplementary Figure 1. A Theory Map of Automatic Imitation.



[1] Detecting the presence of other humans (e.g., Kanwisher, 2010)
[2] Detection of dynamic and goal-directed behaviour (e.g., Caspers et al., 2010)
[3] Domain General control systems are associated with the multiple demand network and by definition operate across multiple tasks / domains (e.g., Duncan, 2010)
[4] Domain Specific control systems are associated with the theory of mind network and are functionally tuned towards social information processing (Brass et al., 2009)
[5] Unintentional copying of others actions in social contexts (see Chartrand & lakin, 2013)

Figure 2. A Theory Map of Automatic Imitation. For more details on the development of Theory Mapping, see Gray (2017); www.theorymaps.org.